Econometric Analysis

Ryuichi Tanaka

Ch. 9 Simultaneous Equations Models

9.1 Autonomy and Causal Effect

For Causality analysis, autonomy of each equation is needed

Autonomous equation : each equation has independent (economic) meaning

Example of Autonomous equations: labor supply function and wage equation

Labor supply function $h^s(w) = \gamma_1 w + \mathbf{z}_1\boldsymbol{\delta}_1 + u_1$ (derived from consumer's utility maximization problem)

Wage equation $w^o(h) = \gamma_2 h + \mathbf{z}_2\boldsymbol{\delta}_2 + u_2$ (derived from firm's profit maximization problem)

Each equation has its own economic meaning (autonomous)

We can discuss causal effect after estimating these equations

Example of non-autonomous model: time allocation between legal activity and crime

Individual i determines the time allocation between legal activity (work) and criminal activity

From FOC of util max prob, $h = \gamma_1 crime + \mathbf{z}_1\boldsymbol{\delta}_1 + u_1$, $crime = \gamma_2 h + \mathbf{z}_2\boldsymbol{\delta}_2 + u_2$

Even if we estimate these equations, we only know the trade-off between these two types of activities, but do not know "an increase in time for legal activity reduces the time for crime " in a causal sense.

If endogenous variables in the system are choice variable of one agent (individual, firm), each equation is not autonomous. In the above example, each equation is autonomous (labor supply by worker, wage setting by firm)

9.2.1 Exclusion Restrictions and Reduced Forms

$$\begin{aligned} y_1 &= \mathbf{y}_{(1)}\boldsymbol{\gamma}_{(1)} + \mathbf{z}_{(1)}\boldsymbol{\delta}_{(1)} + u_1 \\ &\vdots \\ y_G &= \mathbf{y}_{(G)}\boldsymbol{\gamma}_{(G)} + \mathbf{z}_{(G)}\boldsymbol{\delta}_{(G)} + u_G \end{aligned}$$

$\mathbf{y}_{(h)}$: endo$(1\times G_h)$(exclude $\mathbf{y}_h$), $\mathbf{z}_{(h)}$: exo $(1\times M_h)$

Exclusion Restrictions : some variables are not contained in one equation

Ex. labor supply and demand

$h = \gamma_1 w + \delta_1 educ + u_1$, $w = \gamma_2 h + \delta_2 educ + u_2$

We cannot identify the parameters because $\gamma_2 = 1/\gamma_1$, $\delta_2 = -\delta_1/\gamma_1$ (two equations are identical)

Identification by Exclusion Restriction

Labor supply function $h^s(w) = \gamma_1 \log(w) + \mathbf{z}_{(1)}\boldsymbol{\delta}_{(1)} + u_1$

Labor demand function $h^d(w) = \gamma_2 \log(w) + \mathbf{z}_{(2)}\boldsymbol{\delta}_{(2)} + u_2$

In equilibrium, $h^d(w) = h^s(w) = h$**. Setting** $y_1 = h, y_2 = \log(w)$**,**

$y_1 = \gamma_1 y_2 + \mathbf{z}_{(1)}\boldsymbol{\delta}_{(1)} + u_1$ and $y_1 = \gamma_2 y_2 + \mathbf{z}_{(2)}\boldsymbol{\delta}_{(2)} + u_2$

To identify the parameters in the labor supply function, we need to have at least one exogenous variable in $\mathbf{z}_{(2)}$ that is not in $\mathbf{z}_{(1)}$

Reduced Form: sovling these equations for y_2,

$y_2 = \mathbf{z}_{(1)}\boldsymbol{\pi}_{21} + \mathbf{z}_{(2)}\boldsymbol{\pi}_{22} + v_2$

$\boldsymbol{\pi}_{21} = \boldsymbol{\delta}_{(1)}/(\gamma_2 - \gamma_1)$, $\boldsymbol{\pi}_{22} = -\boldsymbol{\delta}_{(2)}/(\gamma_2 - \gamma_1)$, $v_2 = (u_1 - u_2)/(\gamma_2 - \gamma_1)$

Requirement for identification: $\mathbf{z}_{(2)}$ **need to contain at least one exogenous variable that is not in $\mathbf{z}_{(1)}$ and has non zero coefficient in the reduced form**

Example 9.1 labor supply of married women

supply $h^s(w) = \gamma_1 \log(w) + \delta_{10} + \delta_{11} educ + \delta_{12} age + \delta_{13} kids + \delta_{14} othinc + u_1$

demand $h^d(w) = \gamma_2 \log(w) + \delta_{20} + \delta_{21} educ + \delta_{22} exper + u_2$

Under the assumption that the past work experience does not affect labor supply directly, *exper* is excluded from the supply function. Since *exper* is excluded from the supply function, the supply function is identified.

Labor demand function is also identified (why?)

In general necessary condition for identification

Theorem 9.1: In a linear system of equations with exclusion restrictions, a *necessary* condition for identifying any particular equation is that the number of excluded exogenous variables from the equation must be at least as large as the number of included right-hand-side endogenous variables in the equation.

In Example 9.1, the number of endogenous variable in the supply function is one. Since we had one exogenous variable excluded from the supply function, the necessary condition was satisfied

9.2.2 General Linear Restrictions and Structural Equations

$$\mathbf{y}\boldsymbol{\gamma}_1 + \mathbf{z}\boldsymbol{\delta}_1 + u_1 = 0$$
$$\vdots$$
$$\mathbf{y}\boldsymbol{\gamma}_G + \mathbf{z}\boldsymbol{\delta}_G + u_G = 0$$

$\mathbf{y} = (y_1, \ldots, y_G)$ **: all endogenous variables,** $\mathbf{z} = (z_1, \ldots, z_M)$ **: all exogenous variables**

In matrix form: $\mathbf{y\Gamma} + \mathbf{z\Delta} + \mathbf{u} = \mathbf{0}$

For the existence of a reduced form, assume that Γ **is nonsingular**

$$\mathbf{y} = \mathbf{z}(-\mathbf{\Delta\Gamma}^{-1}) + \mathbf{u}(-\mathbf{\Gamma}^{-1}) \equiv \mathbf{z\Pi} + \mathbf{v}, \quad \mathbf{\Lambda} \equiv E(\mathbf{v'v}) = \mathbf{\Gamma}^{-1\prime}\mathbf{\Sigma\Gamma}^{-1}, \quad \mathbf{\Sigma} \equiv E(\mathbf{u'u})$$

<u>Under which conditions, are the parameters in the structural equation identified?</u>

$$\mathbf{y}\boldsymbol{\gamma}_1 + \mathbf{z}\boldsymbol{\delta}_1 + u_1 = 0$$
$$\vdots$$
$$\mathbf{y}\boldsymbol{\gamma}_G + \mathbf{z}\boldsymbol{\delta}_G + u_G = 0$$

wanna identify $\boldsymbol{\beta}_1 \equiv (\boldsymbol{\gamma}_1', \boldsymbol{\delta}_1')'$ in the eq.(1) $\mathbf{y}\boldsymbol{\gamma}_1 + \mathbf{z}\boldsymbol{\delta}_1 + u_1 = 0$**!**

<u>Normalization restriction (NR)</u> one element in γ_1 **must be -1**

<u>Homogeneous linear restrictions (HLR)</u> $\boldsymbol{\beta}_1 \equiv (\boldsymbol{\gamma}_1', \boldsymbol{\delta}_1')'$ has a relation $\mathbf{R}_1\boldsymbol{\beta}_1 = \mathbf{0}$

<u>Theorem 9.3</u> (Order Condition): A necessary condition for identification is $J_1 \geq G-1$ where J_1 is the row dimension of $\mathbf{R}_1$.

<u>Theorem 9.2</u> (Rank Condition): Assume (NR) and (HLR). Then $\boldsymbol{\beta}_1$ is identified if and only if rank $\mathbf{R}_1\mathbf{B} = G-1$ where $\mathbf{B} \equiv \begin{pmatrix} \mathbf{\Gamma} \\ \mathbf{\Delta} \end{pmatrix}$. (solution of $\mathbf{R}_1\boldsymbol{\beta}_1 = \mathbf{0}$ **is unique**)

<u>In sum：how to check if the eq(1) is identifiable</u>

１. **Check Normalization restriction**

２. **Construct the matrix of homogeneous linear restriction** $\mathbf{R}_1$ **for** $\boldsymbol{\beta}_1$

３. **Check order condition** $J_1 \geq G-1$

４. **If rank** $\mathbf{R}_1\mathbf{B} = G-1$**, identified**

Example 9.3 $y_1 = \gamma_{12}y_2 + \gamma_{13}y_3 + \delta_{11}z_1 + \delta_{13}z_3 + u_1,\ y_2 = \gamma_{21}y_1 + \delta_{21}z_1 + u_2,\ y_3 = \delta_{31}z_1 + \delta_{32}z_2 + \delta_{33}z_3 + \delta_{34}z_4 + u_3$

１. y_1 **is on the LHS** $\gamma_{11} = -1$ **(OK)**

２. $\boldsymbol{\beta}_1 = (-1, \gamma_{12}, \gamma_{13}, \delta_{11}, \delta_{12}, \delta_{13}, \delta_{14})'$**,** $\mathbf{R}_1 = \begin{pmatrix} 0 & 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 \end{pmatrix}$**,** $J_1 = 2$

３. **Since** $J_1 = 2$ **and** $G-1 = 2$**, order condition is satisfied. (G=3, M=4)**

４. **Since** $rank\mathbf{R}_1\mathbf{B} = rank\begin{pmatrix} 0 & 0 & \delta_{32} \\ 0 & 0 & \delta_{34} \end{pmatrix} = 1$**, rank condition is not satisfied (unidentified)**

9.3.1 Robustness-Efficiency Trade-off in the system estimation

If the system satisfies rank condition, we can use 3SLS, GMM or just use 2SLS only for the equation in interest. Tests are valid as well.

<u>Merit and demerit of System equation estimation</u>

merit if the specification is correct, system estimation is more efficient than 2SLS equation by equation

Demerit if there is a specification error in anywhere in the system, 3SLS and GMM are inconsistent

Example 9.5, PS 9.9.a

9.4.1 Using Cross Equation Restrictions to Achieve Identification

Identification using the cross equation restriction in the parameters

（１） $y_1 = \gamma_{12} y_2 + \delta_{11} z_1 + \delta_{12} z_2 + \delta_{13} z_3 + u_1$

（２） $y_2 = \gamma_{21} y_1 + \delta_{21} z_1 + \delta_{22} z_2 + u_2$

The eq. (1) is unidentified. The eq. (2) is just identified iff $\delta_{13} \neq 0$

Suppose $\delta_{12} = \delta_{22}$ **(from theory). Then after obtaining** $\hat{\delta}_{22}$ **from the estimation of (2), the eq. (1) becomes (3)** $y_1 - \hat{\delta}_{12} z_2 = \gamma_{12} y_2 + \delta_{11} z_1 + \delta_{13} z_3 + u_1$**.**

We can estimate the eq. (3) by 2SLS with (z_2, z_1, z_3) **as IVs**

9.4.2 Using Covariance Restrictions to Achieve Identification

（１） $y_1 = \gamma_{12} y_2 + \delta_{11} z_1 + \delta_{13} z_3 + u_1$

（２） $y_2 = \gamma_{21} y_1 + \delta_{21} z_1 + \delta_{22} z_2 + \delta_{23} z_3 + u_2$

The eq. (1) is just identified iff $\delta_{22} \neq 0$**. The eq. (2) is unidentified.**

Suppose $Cov(u_1, u_2) = E(u_1 u_2) = 0$ **(from theory). The eq. (1) is identified, obtain** $\hat{u}_1 = y_1 - \hat{\gamma}_{12} y_2 - \hat{\delta}_{11} z_1 - \hat{\delta}_{13} z_3$ **from the 2SLS estimation of the eq. (1).**

Since $Cov(u_1, u_2) = E(u_1 u_2) = 0$ **and** $E(y_1 u_1) \neq 0$**, the eq. (2) is identified by 2SLS with** $(z_2, z_1, z_3, \hat{u}_1)$ **as IVs**

(In general, fully recursive system is identified)

9.5 SEMs Nonlinear in Endogenous Variables

9.5.1 Identification

Ex. demand and supply

supply: $\log(q) = \gamma_{12} \log(p) + \gamma_{13} [\log(p)]^2 + \delta_{11} z_1 + u_1$

demand: $\log(q) = \gamma_{22} \log(p) + \delta_{22} z_2 + u_2$ $\qquad E(u_1 \mid \mathbf{z}) = E(u_2 \mid \mathbf{z}) = 0$

setting $y_1 = \log(q), y_2 = \log(p)$, $y_1 = \gamma_{12} y_2 + \gamma_{13} y_2^2 + \delta_{11} z_1 + u_1$, $\qquad y_1 = \gamma_{22} y_2 + \delta_{22} z_2 + u_2$

Solution by Fisher (1965)

1. **Consider nonlinear endogenous variable as a new endogenous variable**
2. **We do not write down the third equation for the new endogenous variable, but treat the new variable as endogenous variable when we consider the rank condition. If the rank condition is satisfied, it is identified**

Ex. consider $y_3 = y_2^2$ **as a new endogenous variable,**

$y_1 = \gamma_{12} y_2 + \gamma_{13} y_3 + \delta_{11} z_1 + u_1$, $\qquad y_1 = \gamma_{22} y_2 + \delta_{22} z_2 + u_2$

As the rank condition for identification, use G=2 (not G=1)

(The above method is identical to considering the third equation explicitly $y_3 = \pi_{31} z_1 + \pi_{32} z_2 + \pi_{33} z_1^2 + \pi_{34} z_2^2 + \pi_{35} z_1 z_2 + v_3$ **and considering rank condition)**

9.5.2 Estimation

The choice of instruments is arbitral, but in general use exogenous variables in the system and higher order terms and cross terms as instruments

For the estimation, use 2SLS, 3SLS, GMM as they are
(it is wrong approach to use fitted value obtained from the first stage (forbidden regression))

Example 9.6

PS 12

I. Example 9.5
II. 9.1
III. 9.3